SecBot: a Business-Driven Conversational Agent for Cybersecurity Planning and Management

Figueredo Franco, Muriel ; Rodrigues, Bruno ; John Scheid, Eder ; Jacobs, Arthur ; Killer, Christian ; Granville, Lisandro Zambenedetti ; Stiller, Burkhard

Abstract: Businesses were moving during the past decades to-ward full digital models, which made companies face new threatsand cyberattacks affecting their services and, consequently, theirprofits. To avoid negative impacts, companies' investments incybersecurity are increasing considerably. However, Small andMedium-sized Enterprises (SMEs) operate on small budgets,minimal technical expertise, and few personnel to address cy-bersecurity threats. In order to address such challenges, it isessential to promote novel approaches that can intuitively presentcybersecurity-related technical information.This paper introduces SecBot, a cybersecurity-driven conver-sational agent (i.e., chatbot) for the support of cybersecurityplanning and management. SecBot applies concepts of neuralnetworks and Natural Language Processing (NLP), to interactand extract information from a conversation. SecBot can(a)identify cyberattacks based on related symptoms,(b)indicatesolutions and configurations according to business demands,and(c)provide insightful information for the decision on cy-bersecurity investments and risks. A formal description hadbeen developed to describe states, transitions, a language, anda Proof-of-Concept (PoC) implementation. A case study and aperformance evaluation were conducted to provide evidence ofthe proposed solution's feasibility and accuracy

DOI: https://doi.org/10.23919/CNSM50824.2020.9269037

SecBot: a Business-Driven Conversational Agent for Cybersecurity Planning and Management

Muriel Figueredo Franco[1], Bruno Rodrigues[1], Eder John Scheid[1], Arthur Jacobs[2],
Christian Killer[1], Lisandro Zambenedetti Granville[2], Burkhard Stiller[1]
[1] *Communication Systems Group CSG, Department of Informatics IfI, University of Zürich UZH*
Binzmühlestrasse 14, CH—8050 Zürich, Switzerland
[2]*Computer Networks Group, Institute of Informatics, Federal University of Rio Grande do Sul UFRGS*
Av. Bento Gonçalves, 9500, Porto Alegre, Brazil
[franco|rodrigues|scheid|killer|stiller]@ifi.uzh.ch, [asjacobs|granville]@inf.ufrgs.br

***Abstract*—Businesses were moving during the past decades toward full digital models, which made companies face new threats and cyberattacks affecting their services and, consequently, their profits. To avoid negative impacts, companies' investments in cybersecurity are increasing considerably. However, Small and Medium-sized Enterprises (SMEs) operate on small budgets, minimal technical expertise, and few personnel to address cybersecurity threats. In order to address such challenges, it is essential to promote novel approaches that can intuitively present cybersecurity-related technical information.**

This paper introduces SecBot, a cybersecurity-driven conversational agent (*i.e.*, chatbot) for the support of cybersecurity planning and management. SecBot applies concepts of neural networks and Natural Language Processing (NLP), to interact and extract information from a conversation. SecBot can *(a)* identify cyberattacks based on related symptoms, *(b)* indicate solutions and configurations according to business demands, and *(c)* provide insightful information for the decision on cybersecurity investments and risks. A formal description had been developed to describe states, transitions, a language, and a Proof-of-Concept (PoC) implementation. A case study and a performance evaluation were conducted to provide evidence of the proposed solution's feasibility and accuracy.

I. Introduction

Businesses becomes proportionally more exposed to cyberattacks as their reliance on Information and Communications Technologies (ICT) increases. As result, companies investments in cybersecurity naturally increase [10]. While large companies such as banks and governmental entities spend significant funds on adopting cybersecurity best practices and training dedicated technical personnel, Small and Medium-sized Enterprises (SMEs) often underinvest and lack efficient strategies to protect their Information Technology (IT) services and value chains they are part of [5]. In addition, SMEs tend to show a misperception of their cybersecurity conditions, as a recent survey reveals [3]. While 60% of US and UK SMEs believe their businesses are unlikely to be targeted by cyberattacks, the reality is the opposite, with a significant amount of breaches and cyberattacks targeting SMEs [30].

The adoption of efficient cybersecurity strategies in SMEs is challenging because of constraints mainly associated with the lack of a cybersecurity budget, unskilled human resources, and limited time allocated to cybersecurity planning [12]. This can lead to disastrous impacts on business, including financial losses due to cyberattacks, mitigation of costs, and inefficient management of protections [23]. From a human-centric perspective, simplifying the cybersecurity decision-making process requires clear and straightforward approaches for SMEs [22]. It is essential to promote novel approaches that present cybersecurity technical information in an intuitive, user-friendly way [19], allowing less-skilled personnel to make informed decisions while maintaining a proper level of protection of their businesses. SMEs can benefit from adopting faster and cheaper cybersecurity strategies, *e.g.*, by minimizing human experts' need while reducing costs by efficiently investing in defense mechanisms.

Conversational agents (*i.e.*, chatbots) [20] have been recently highlighted as an ally to enhance business' cybersecurity adoption by sharing network and security information with non-technical staff [6] [2]. Advances in Natural Language Processing (NLP) [14] — driven by novel Machine Learning (ML) techniques [25] — led to conversational interfaces capable of extracting meaningful information and simplifying interactions between humans and machines. Compared to, *e.g.*, command-lines and technical dashboards, chatbots *(i)* provide a straightforward interaction using natural language, *(ii)* enable faster decision-making, and *(iii)* speed-up complex processes. The Cyber Helpline chatbot in the UK [24] was proposed to provide immediate advice to citizens on how to deal with cybersecurity issues. However, even with those benefits, the employment of chatbots in the context of SME cybersecurity is still scarce and limited to very specific scenarios. Hence, the current state-of-the-art neither fully covers the demands of SMEs nor considers barriers for cybersecurity adoption in SMEs (*e.g.*, awareness of standards, limited internal knowledge, and lack of clear implementation guidelines) [7].

In this context, SecBot, a cybersecurity-driven conversational agent, is introduced here to help non-expert users take informed and efficient cybersecurity decisions, reducing the risk of economic impacts due to business disruptions. For that, SecBot is designed to interact with non-experts to extract information on cybersecurity demands and business requirements. SecBot is able to *(i)* understand symptoms and

978-3-903176-31-7

business risks to correlate with potential cyberattacks, helping users comprehend incidents and their impacts, *(ii)* provide recommendations for actions in different levels of abstraction, such as which efforts are required to avoid or to mitigate problems, and *(iii)* support the configuration (*e.g.*, in-house firewall) or acquisition of protections, preparing actions (*e.g.*, command-lines or configuration files) required to configure or deploy a solution. The feasibility of SecBot is evaluated by conducting a case study and by analyzing its performance.

The remainder of this paper is organized as follows. Related work on chatbots is reviewed in Section II. The SecBot solution is introduced in Section III, where design details are provided. Section IV provides an evaluation of SecBot's performance, including a case study and discussion of achievements and limitations. Finally, Section V draws conclusions and comments on future work.

II. Related Work

Conversational agents have been widely used in a variety of areas and different ways. A survey on enabling technologies and application scenarios is presented by [1], providing an overview and comparison on various Natural Language Processing (NLP) techniques and outlining significant factors that impact the design of a chatbot. While NLP techniques vary according to input data, the authors recommend limiting the scope of a chatbot to avoid general-purpose agents that often require more comprehensive knowledge bases. Thus, different fields can benefit from chatbots trained and designed for specific purposes, such as self-driving network management [13]). As the extraction of the correct information from a conversation is critical, [34] surveys recent advances in named entity recognition by using machine learning models, which shows, *e.g.*, that neural networks models outperform other models to recognize entities.

[33] presents a generic chatbot model to answer customer requests based on social media interactions. The authors employ Deep Learning (DL) techniques based on Long Short-Term Memory (LSTM) networks to generate responses for customer-service requests on social media. In contrast to [1], which limits the learning scope, [33] uses a comprehensive learning base relying on using social media (mainly Twitter). An example of an extensive application of chatbots is the Microsoft XiaoIce [37]. As of today, XiaoIce is covering 660 million users across five different countries, averaging 23 conversations per second, gathering data from multiple social networks (mainly from China). This tool provides excellent lessons for the development, improvement, and application of chatbots. Its learning algorithms have been used to infer semantics in massive amounts of data to provide personal assistants' emotional connection. Other relevant applications are seen in [36] and [16], in which the usage of chatbots is proposed as a hook for identifying criminals on the Internet and as an ally to cyberpedophile identification.

Specifically related to information security, [11] investigated chatbots as a tool for IT Security Training, providing hints and elementary training steps concerning the handling of passwords, privacy, and secure Internet browsing. This tool was planned to be used in large companies, where employees' face-to-face training is infeasible. [18] proposed a conversational agent to address the complexity of how to present network information to non-technical users about the behavior of IoT devices, helping identify when devices are part of a botnet. However, none of these solutions focus on business demands or directly explore different tasks involved in the decision-making, configuration, and cybersecurity management.

Even though exist clear indications of benefits of exploiting AI-based chatbots for the cybersecurity field (*e.g.*, by simplifying the access to the security information for different stakeholders [6]), there is still a lack of work exploring that for a cybersecurity adoption and management. Therefore, SecBot is designed to address this gap, mainly focusing on the demands of the SME sector, but designed in an extensible way to covers other IT sectors too.

III. SecBot Design

Two fundamental concepts are required for conversational agents: *Intents* and *Entities*. These concepts determine the basis to describe information and flows supported by SecBot. *Intents* refer to user's intentions when interacting with the chatbot, and *Entities* are defined to extract specific terms or values. Extracting entities and intent classification typically involves an ML architecture. While non-ML approaches do exist [14], they are normally outperformed by supervised learning algorithms [34], which can generalize the information extraction process by understanding the context of input phrases. In the case of SecBot, a Dual Intent and Entity Transformer (DIET) [4] architecture is used for intent classification and entity extraction, implemented by the Rasa framework [21]. The DIET classifier relies on a transformer neural network [29] to encode input text with context, Conditional Random Fields (CRFs) [15] to identify and extract entities from text encoded, and dot-product similarity [31] to classify the input intent.

While *Intents* (*cf.* Table I) identify users that want to find protection according to the budget available or want to ask for help to configure efficient protection, *Entities* are used to extract specific terms or values (*cf.* Table II) from the user intent to provide a correct response. To reach accurate responses, all entities are connected to knowledge databases, which describe values accepted for each of the specific entities. About 150 entries are defined for *Entities* (*cf.* Table II) of SecBot. New entries for these *Entities* as well as new *Intents* can be added, such that the SecBot can cover different scenarios and demands.

After identifying the user's intent and extracting input entities from the input text, the SecBot needs to decide upon which action to take to best help the user. To that end, another important concept for conversational agents needs to be defined: *Stories*. A single *Story* defines those steps SecBot can take in response to a user's input, resulting in multiple possible conversation flows. For example, after recognizing

the intent *attack_notification* and if the next one is the Intent *attack_details*, a message is sent asking for the budget available to invest in protection, before issuing a recommendation. However, if the next intent recognized is *problem_desc*, a different action will be executed to identify the type of attack. Thus, the definition of *Stories* is critical, given that it is used to train the solution to recognize the context of a conversation and to select the next actions or flows.

TABLE I: Examples of Intents Implemented by the SecBot

Intent	Example	Associated Entities
attack_notification	My Windows systems are under a Ransomware attack	@target, @attack_name
attack_details	It is a WannaCry attack	@attack_type
target	The target is my database	@target
problem_desc	My server is receiving a lot of requests from different IPs	@symptom, @target
solution_config	I want to block an SYN flood using my IPTables	@solution, @technology, @target, @attack_name, @action
solution_support	How can I block a specific port using UFW?	@operator, @object, @solution
rosi_calc	Should I invest in backups against Ransomware impacts?	@attack_name
critical_data	I have almost 10 TB of critical data	@cardinal

SecBot supports functions that can be run as an action in response to users' inputs, according to an identified *Intent*, such as providing feedback messages, running arbitrary code (*i.e.*, custom actions), or listening for new inputs. Based on that, SecBot implements different custom actions (*cf.* Sections III-A, B, and C) that run actions according to different scenario flows. These custom actions involve *(i)* finding the best solution for a request, *(ii)* identifying the type of attack based on symptoms, *(iii)* helping during the configurations of in-house protections, and *(iv)* calculating metrics related to economic impacts of different cyberattacks.

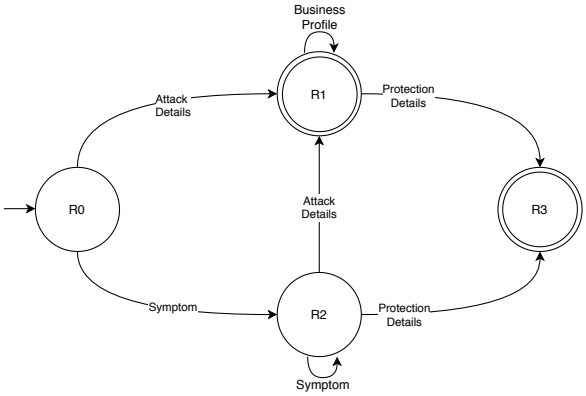
TABLE II: Examples of Entities Supported by the SecBot

Entity	Description	Input's Example
@attack_name	Name of the attack	I am being target of a @DDoS Attack.
@attack_type	Type of attack	It looks like a @SYN flood.
@target	Target of the attack or the component with symptoms	The target is my @Windows systems. It is my @database server.
@symptom	Describe specific problems or symptoms	My server is receiving @a lot of requests.
@budget	Amount and currency available to invest	My budget is @5000 EUR.
@solution, @technology	Describe in-house solutions	I have an @IPtables running on @Linux.
@operator	Describes the users' required action	I want help to @block an IP traffic using the UFW firewall.
@object	Explicitly describes an element to apply the operator	I want to block the @Port 22 using IPtables.

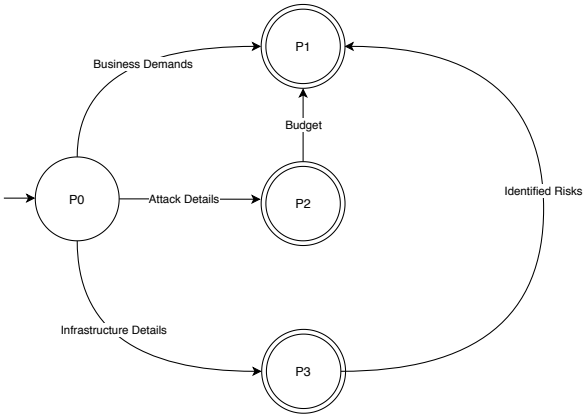
During the training phase of the SecBot, besides database entries and *Intents*, different *Stories* have to be defined for the supervised learning to allow the implemented Rasa neural network algorithm to obtain sufficient knowledge to extract and process information. Thus, it is possible to determine which action to take next during a conversation correctly. These *Stories* were defined to cover SecBot scenarios, being able to predict a correct flow based on an identified *Intent*.

A. SecBot Scenarios

Two approaches are defined to describe different scenarios and to guide users during the interaction with the SecBot: the Reactive `R` and the Proactive `P` approaches. These approaches define, respectively, situations where the user wants to react to protect against an imminent attack or a user that wants to operate a better plan defining the business cybersecurity strategy. These two approaches are divided into six different flows that can be combined to provide a more accurate and complete answer to the user.



(a) Reactive (R) Flow



(b) Proactive (P) Flow

Fig. 1: Finite Automaton for the SecBot Scenarios

Figure 1 (a) describes the finite automaton for **reactive** scenarios. R_1 represents a conversation, where the user knows technical details of the attack (*e.g.*, type of attack or log files) and wants to know which solution matches his/her budget and demands. R_2 focuses on understanding symptoms associated with cyberattacks and problems, thus helping users find a suitable solution. Lastly, the flow resulting in the final state R_3 covers users that already deployed protection solutions, but need help to configure these.

The finite automaton for **proactive** scenarios is presented in Figure 1 (b). P_1 assumes users who want to reduce the economic impacts of threats in their business. Different metrics can be employed to provide useful information, directly helping during the decision related to where and when investing in cybersecurity. *E.g.*, the Return On Security Investment (ROSI) metric [27] is calculated using the user's inputs and business

requirements to provide insights about whether to contract a solution, assume risks, or even acquire a cybersecurity insurance coverage. Furthermore, based on its knowledge database, the agent can suggest actions to reduce costs and to avoid a financial loss for specific business sectors. Scenario P_2 covers the conversation flow in which users want to proactively protect their systems against specific cyberattacks (*e.g.*, WannaCry Ransomware or Mirai Botnet). For that, recommendations for updates, configurations, or solutions to be acquired can be provided. Finally, P_3 considers requests about the most common risks and vulnerabilities according to the business configuration, sector, and information provided.

A business profile descriptor, based on a JSON structure as defined in previous work [8], can be configured by users to provide the SecBot with a detailed view about their business. This information is used for the recommendation process and steps requiring specific information on the business organization (*e.g.*, number of employees, regulations, sector, or underlying security configurations/demands). To choose the best solution from a list of possible protections, the SecBot is integrated with MENTOR, a recommender system for the protection of services [8].

Different custom actions are presented next to handle information obtained during the conversation, providing accurate answers for specific cases, where algorithms and calculations are required to process the output, such as those specific reactive and proactive flows described. Custom actions are provided to SecBot to *(a)* identify a cyberattack based on a list of presented problems or symptoms, *(b)* provide configurations for protections according to requests, and *(c)* conduct an economic analysis based on user's requests to support the decision-making.

B. Attack Identification

The symptoms or problems extracted from the conversation can be used to identify the attack described by the user. To that end, a decision-tree containing the relationship between known attacks and associated symptoms is proposed as a custom action, which receives a list of symptoms and returns the related attack for the user. This action is directly related to the intent named as *problem desc* (*cf.* Table I), which is recognized when the user describes problems without a technical understanding about what is happening.

Figure 2 shows an example of the attack tree structure. The SecBot starts with an initial tree containing examples of well-known attacks (*e.g.*, Distributed Denial-of-Service - DDoS and ransomware) relationships and their symptoms. Thus, the user's described symptoms are checked in the attack tree. If the resulting path ends in a leaf, it means that the attack was identified. Thus, using a *Server* as a target, the symptoms *"receiving many requests"* and *"many of them are SYN packets"* can result in the identification of an SYN flood attack. The same approach can be applied for different attacks in which previously known symptoms can be used to create the attack decision-tree. If the path cannot achieve a leaf, it means that the attack cannot be identified, resulting in negative feedback sent to the user.

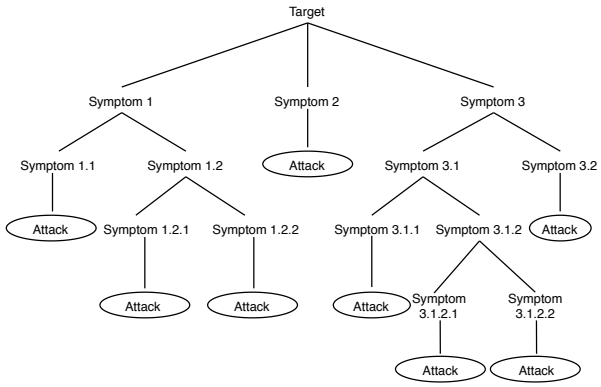


Fig. 2: Symptoms' Tree Structure to Search for an Attack

C. Protection Configuration

The SecBot also interprets requests for help to configure protection already available in-house. Hence, entities are extracted to understand *(i)* the intent of the user, which includes the name of the solution available, *(ii)* the operator (*e.g.*, block, allow, or protect), and *(iii)* the attack type for which the user wants a specific configuration. Based on these entities, the SecBot can determine the associated configuration or provide the syntax for the user to create his/her own configuration.

```
<input>: "I have an IPtables installed and I want
    to protect my network against ICMP flood"
Entities_Extraction {
    "intent": solution_configuration
    "solution": IPtables
    "operator": protect
    "target": network
    "attack_name": ICMP flood
}
<custom_action>: find_configuration(solution,
    action, target, attack_name)
<output>: "The command for your configuration
    request is: iptables -t mangle -A PREROUTING -
    p icmp -j DROP"
```

Listing 1: Example of SecBot Processing and Output Based on a User's Input

Listing 1 presents the input and output for scenarios where users want to protect the network from an imminent attack (*i.e.*, reactive) or anticipate (*i.e.*, proactive) this type of attack to avoid damages. *E.g.*, the request *"I have an IPtables installed and want to protect my network against ICMP flood"* results in a message containing a configuration for *protection* against *ICMP flood* tailored for the IPTables packet filtering solution. This configuration is provided as a JSON structure stored by the SecBot, which maps different solutions, configurations, and commands.

D. Cybersecurity Investment

The extraction of entities related to the attack (*e.g.*, @*attack_name* and @*attack_type*) and to the business itself (*e.g.*,

budget, sector, amount of critical services, and data) is essential for the conversational agent to understand the scenario and to achieve accurate information to calculate the ROSI metric. This metric is defined by Equation 1, where the Reactive Mitigation Cost (RMC) and the total cost (*i.e.*, financial impacts of risk exposure) of a specific attack are calculated given a time-frame ΔT. Furthermore, the Proactive Mitigation Cost (PMC) is used for the ROSI calculation, which defines the cost of investing in approaches or solutions to anticipate threats and avoid future damage (*e.g.*, financial loss). Thus, the higher ROSI is, the more the business is recommended to follow a proactive approach (*e.g.*, to contract backups services or pay for a continuous cloud-based DDoS protection). Otherwise, if ROSI's result is near 0, the business can, *e.g.*, assume risks of economic impacts regarding a possible threat or specific cyberattack.

$$ROSI = \Delta T * \frac{(T_{costs} * RMC) - PMC}{PMC} \quad (1)$$

During the conversation flow, the SecBot can map the attack type based on a specific structure, associating attacks to possible proactive approaches. *E.g.*, for a Ransomware attack, information about the amount of data available (in GB) is required to measure costs of *(i)* maintaining a full backup to recover from an attack or *(ii)* a cybersecurity insurance. This information is crucial to calculate the ROSI based on this type of attack's possible financial losses. Also, if the user is not able to provide details about specific backup prices for calculation, the SecBot uses an internal database with average costs for different services (*e.g.*, backup, DDoS protection, and anti-phishing) and different attacks (*e.g.*, rescue price for a Ransomware and costs per hour of a DDoS) to provide an approximated ROSI, even with missing inputs from users.

IV. SecBot Evaluation

To evaluate the SecBot, a Proof-of-Concept (PoC) was developed and evaluated using Rasa 2.0rc2 [21], an open-source machine learning framework to build contextual AI agents and chatbots. SecBot's code and training data set are publicly available [9]. The implemented solution relies on the Rasa framework abstractions of the underlying NLP and ML algorithms to simplify the design and handling of Entities and Intents. Custom actions were developed using Python 3.8.3, while the knowledge databases are described as plain text or JSON files. The evaluation was performed using a Dell XPS desktop with the configuration of an Intel Core i7-3770 at 3.40 GHz, 32 GByte of RAM, running a Linux Ubuntu 18.04 LTS 64-bit with the Linux Kernel version 5.3.0-53.

The current training of SecBot is done using a neural network implemented in Rasa to select the next action, which is described as an LSTM architecture defined in [32]. For the training of the neural network, it receives the user's phrase as input and actions as output. During the training phase, it is used as a fitting model with 958 samples (*i.e.*, examples of intents and entities) and a validating split of 0.1 (*i.e.*, 10% of the training dataset as validation data only), which covers 15 different conversation flows with 100% of accuracy for the intent and entities extraction. These results indicate that SecBot can map the conversation for the correct intent available, thus, also being able to extract entities.

In terms of scalability, a stress test revealed that one single instance of SecBot can handle 20 messages per second. Among the currently supported custom actions, a more time-consuming request is the one to identify an attack, using symptoms in the attack tree, which have a computational complexity of $O(nlogn)$. In a simulation with an attack tree containing 100 symptoms and 30 attacks (*i.e.*, leaves), the time for the SecBot to process the request and return the correct attack is less than 2 s on average, considering 1,000 repetitions.

A. Case Study

The case study was conducted by interacting with an instance of SecBot's prototype running on Telegram, a popular messenger platform [28]. The application interface provided by Telegram simplifies the process of presenting interactions of the business and the SecBot, thus offering a better usability and user-acceptance. However, it is possible to conduct the same case study using the terminal provided by the Rasa framework or even integrating it with other messenger platforms. It is assumed that an SME faces problems in its server infrastructure and wants to find a solution to solve this issue initially, followed by the configuration of on-site protection (*i.e.*, IPtables) and the calculation of ROSI for investments to reduce impacts of a possible ransomware attack.

Users start a chat with the SecBot and ask for help. Symptoms include a server overload with many requests from many different IP addresses, which is initially identified as a DDoS attack. After more symptoms are described and by searching the attack tree (*cf.* Figure 2), the cyberattack is recognized as a DDoS attack characterized by different hosts sending a flood of SYN requests. Based on this information, the user can ask for protection to help against the attack. The user is asked about his/her budget available to invest in protection. Thus, by using details provided in the business profile (*e.g.*, regulations, region, and business sector), the SecBot can select and recommend, from a list of protections against SYN floods, which protection suits best user demands and budget available.

The user continues the conversation for proactively addressing other aspects that can impact the business. This proactive scenario and its interactions sees the user asking to support the blocking of port scanning on his/her network. If business protections are not described in the business profile configurations, the SecBot asks whether the user already has a solution installed. In this case, IPtables is available running on the business infrastructure. The SecBot can check in its protection configuration descriptor the correct configuration, and then the proper command is provided for the user to block port scanning. Finally, the user checks with the SecBot about the benefits of investing in backups as a proactive approach to reduce impacts of ransomware attacks, since it can cause all critical business files to be encrypted, requiring rescue for the decryption key. This type of attack typically results in business

disruption, financial loss, and also reputation harm. To provide an answer to such a request, the SecBot checks the business profile to understand how much critical data the business has and what the business revenue is. This information is provided in a JSON file used as a descriptor (*i.e.*, business profile) for business configurations and the organization, which can be used as inputs on demand. The downtime average for the business with similar characteristics (*e.g.*, sector and amount of data) is considered for the analysis, too. Based on all this information, the ROSI (*cf.* Equation 1) is calculated and provided to the user, followed by a final recommendation, which in this case, means that an investment in backups is recommended.

Based on this case study, it is possible to observe the feasibility of the SecBot by providing interactions that cover different flows of the conversation to help in relevant cybersecurity-related tasks. These scenarios encompass the support to react against a cyberattack, configure and manage an existent solution according to the business goals, and obtain information for an efficient cybersecurity planning. Also, the performance of the SecBot is highlighted by answering requests and correctly extracting the information required for these scenarios.

B. Discussion

The SecBot shows opportunities to simplify the different steps involved in cybersecurity management. Challenges to chatbots are also highlighted, since the accuracy achieved by supervised learning methods is directly related to the quality of inputs used. For these scenarios and flows defined, the accuracy of answers provided was precise and useful to address users' demands. The current state, as observed in the PoC implemented, provides directions and shows the benefits of addressing cybersecurity-related information using conversational agents. Custom actions, developed as contributions of this work, indicate the path for further implementations and highlight the proposed solution's extensibility.

Given that the SecBot's prototype has been evaluated by using selected information and scenarios, it is possible to learn new information for handling more requests and conversation flows. There are opportunities to improve the training phase by creating new *Stories* and considering different datasets available for cybersecurity [35], such as describing more attack characteristics and their relationships. By building a larger dataset of cybersecurity-related information, it is possible to define additional *Entities* to extract from a conversation, thus, resulting in different flows and scenarios covered. In the same way, new *Intents* and scenarios can be defined based on the amount of information that the SecBot can extract. Such *Intents* need to be defined considering the actual demands of businesses, thus resulting in different custom actions to be implemented to address specific requirements.

In terms of scalability, several instances of the SecBot can be provided quickly in order to address high demands for interactions. As one instance can handle 20 messages per second, it is reasonable to assume that a single instance of the SecBot can be used by many businesses simultaneously, such as processing more than 100 scenarios (equalling the case study as presented) in one minute. Thus, despite relying on similar underlying data sources, each instance runs independently from the others in a modular fashion via replication. In terms of security, it is an option that each SME can run locally their own instance of the chatbot, which increases the means to operate on dedicated resources in a controlled environment, also allowing to have a knowledge database customized according to the specific demands of that business. It also can scale to complex problems and solutions. However, it depends how to define the correct training data set to use to avoid an over-fitting of the machine learning model being used, *i.e.*, ensuring that the model will be able to extrapolate the knowledge of complex scenarios and not only perform with trained scenarios.

Although this work on the SecBot is motivated by identifying the benefits and challenges of chatbots for SMEs, large companies can also benefit. Professionals with prior knowledge in cybersecurity can explore this approach to meet different goals. Cybersecurity analysts can interact with the SecBot to find a fast and accurate answer for a customer request regarding technical and economic aspects related to SMEs' cybersecurity. Also, mechanisms can be implemented to help large companies justify their investments on a specific solution or cybersecurity strategy, such as understanding requirements to define directions of their bug bounties programs. This can help build foundations for long-term cybersecurity strategies rather than sporadic engagements of specialists [17]. Another use comprises the opportunity offered to cybersecurity companies to develop their own solutions to be integrated with the SecBot.

V. Conclusions and Future Work

In conclusion and due to the fact that SecBot combines the description of a formal language with Machine Learning (ML) and state-of-the-art aspects of cybersecurity, the chatbot introduced interacts with users and provides information according to their requests and demands. Since the SecBot executes custom actions to find the best protection configuration for a business, the SecBot can identify attacks during the conversation or can provide insights about risks and economic impacts of possible cyberattacks. Due to language and respective different scenarios based on the finite automata theory and the use of extensible databases, the SecBot does implement custom actions, which can be extended to cover further scenarios and additional demands to support cybersecurity in SMEs and other companies.

In terms of future work, different approaches and ML techniques, such as deep reinforcement learning [26], can be explored to maintain a higher accuracy, when the number of flows accepted and its complexity increases considerably. Finally, extensive evaluations with real-users in different conversation flows are planned for, going well beyond the coherent and full technical design and functional evaluation of the approach itself as outlined in this paper.

ACKNOWLEDGEMENTS

This paper was supported partially by *(a)* the University of Zürich UZH, Switzerland and *(b)* the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No. 830927, the CONCORDIA Project.

REFERENCES

[1] S. A. Abdul-Kader and J. Woods, "Survey on Chatbot Design Techniques in Speech Conversation Systems," *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 7, 2015.

[2] Bobby Filar, "Artemis: an Intelligent Assistant for Cyber Defense," 2017, https://www.elastic.co/blog/artemis-intelligent-assistant-cyber-defense, last visit July 29, 2020.

[3] Bullguard, "New Study Reveals One In Three SMBs Use Free Consumer Cybersecurity And One In Five Use No Endpoint Security At All," February 2020, https://www.bullguard.com/press/press-releases/2020/new-study-reveals-one-in-three-smbs-use-free-consu.aspx, last visit July 20, 2020.

[4] T. Bunk, D. Varshneya, V. Vlasov, and A. Nichol, "DIET: Lightweight Language Understanding for Dialogue Systems," 2020, https://arxiv.org/abs/2004.09936, last visit July 1, 2020.

[5] Capgemini Invent, European Digital SME Alliance, and Executive Agency for Small and Medium-sized Enterprises (European Commission), Technopolis, "Skills for SMEs: Cybersecurity, Internet of things and Big Data for Small and Medium-sized Enterprise," December 2019, https://op.europa.eu/en/publication-detail/-/publication/82aa7f66-67fd-11ea-b735-01aa75ed71a1/language-en, last visit July 20, 2020.

[6] Dania Ben Peretz, "A Siri for Network Security: How Chatbots can Enhance Business Agility," 2020, https://www.infosecurity-magazine.com/opinions/network-chatbots-agility/, last visit July 29, 2020.

[7] European Union Agency for Network and Information Security (ENISA), "Information Security and Privacy Standards for SMEs," June 2016, https://www.enisa.europa.eu/publications/standardisation-for-smes, last visit May 2, 2020.

[8] M. Franco, B. Rodrigues, and B. Stiller, "MENTOR: The Design and Evaluation of a Protection Services Recommender System," in *15th International Conference on Network and Service Management (CNSM 2019)*. Halifax, Canada: IEEE, October 2019, pp. 1–7.

[9] M. Franco, "SecBot Implementation," June 2020, https://gitlab.ifi.uzh.ch/franco/secbot, last visit July 10, 2020.

[10] Gartner Research, "Forecast Analysis: Information Security, Worldwide, 2Q18 Update," 2018, https://www.gartner.com/en/documents/3889055, last visit March 15, 2020.

[11] I. Gulenko, "Chatbot for IT Security Training: Using Motivational Interviewing to Improve Security Behaviour," in *International Conference on Analysis of Images, Social Networks and Texts (AISN)*, Yekaterinburg, Russia, July 2014, pp. 1–10.

[12] M. Heikkilä, A. Rättyä, S. Pieskä, and J. Jämsä, "Security Challenges in Small- and Medium-sized Manufacturing Enterprises," in *International Symposium on Small-scale Intelligent Manufacturing Systems (SIMS)*, Narvik, Norway, June 2016, pp. 25–30.

[13] A. Jacobs, R. Pfitscher, R. Ribeiro, R. Ferreira, L. Granville, and S. Rao, "Deploying Natural Language Intents with Lumi," in *ACM SIGCOMM 2019 Conference Posters and Demos*, Beijing, China, August 2019, pp. 82–84.

[14] D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2019.

[15] J. D. Lafferty., A. McCallum, and F. C. N. Pereira, "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data," in *18th International Conference on Machine Learning (ICML)*, San Francisco, CA, USA, June 2001, pp. 282–289.

[17] S. S. Malladi and H. C. Subramanian, "Bug Bounty Programs for Cybersecurity: Practices, Issues, and Recommendations," *IEEE Software*, vol. 37, no. 1, pp. 31–39, 2020.

[16] C. Laorden, P. Galán-García, I. Santos, B. Sanz, J. H. Gomez, and P. Bringas, "Negobot: A Conversational Agent Based on Game Theory for the Detection of Paedophile Behaviour," *Advances in Intelligent Systems and Computing*, vol. 189, pp. 261–270, January 2013.

[18] C. D. McDermott, B. Jeannelle, and J. P. Isaacs, "Towards a Conversational Agent for Threat Detection in the Internet of Things," in *International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, Mississippi, USA, June 2019, pp. 1–8.

[19] C. Paulsen, "Cybersecuring Small Businesses," *IEEE Computer*, vol. 49, no. 8, pp. 92–97, 2016.

[20] A. M. Rahman, A. A. Mamun, and A. Islam, "Programming Challenges of Chatbot: Current and Future Prospective," in *IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, Dhaka, Bangladesh, December 2017, pp. 75–78.

[21] Rasa Technologies, "Rasa: Open Source Conversational AI," June 2020, https://rasa.com/, last visit July 10, 2020.

[22] K. Renaud and G. R. S. Weir, "Cybersecurity and the unbearability of uncertainty," in *Cybersecurity and Cyberforensics Conference (CCC 2016)*, Amman, Jordan, August 2016, pp. 137–143.

[23] B. Rodrigues, M. F. Franco, G. Paranghi, and B. Stiller, "SEConomy: A Framework for the Economic Assessment of Cybersecurity ," in *16th International Conference on the Economics of Grids, Clouds, Systems, and Services (GECON 2019)*. Leeds, UK: Springer LNCS, September 2019, pp. 1–9.

[24] Rory Innes, "The Cyber Helpline - Expert Advice For Victims of Cybercrime," 2019, https://www.thecyberhelpline.com/, last visit July 29, 2020.

[25] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. USA: Pearson, 2020.

[26] I. V. Serban, C. Sankar, M. Germain, S. Zhang, Z. Lin, S. Subramanian, T. Kim, M. Pieper, S. Chandar, N. R. Ke *et al.*, "A Deep Reinforcement Learning Chatbot," 2017, https://arxiv.org/abs/1709.02349, last visit July 12, 2020.

[27] W. Sonnenreich, J. Albanese, B. Stout *et al.*, "Return On Security Investment (ROSI) - A Practical Quantitative Model," *Journal of Research and practice in Information Technology*, vol. 38, pp. 45–52, 2006.

[28] Telegram, "Telegram Messenger - A New Era of Messaging," 2020, https://core.telegram.org/api, last visit July 15, 2020.

[29] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention Is All You Need," 2017, https://arxiv.org/abs/1706.03762, last visit July 1, 2020.

[30] Verizon, "2019 Data Breach Investigations Report," February 2020, https://enterprise.verizon.com/resources/reports/dbir/, last visit July 20, 2020.

[31] V. Vlasov, J. E. M. Mosig, and A. Nichol, "Dialogue Transformers," 2019, https://arxiv.org/abs/1910.00486, last visit July 5, 2020.

[32] V. Vlasov, A. Drissner-Schmid, and A. Nichol, "Few-Shot Generalization Across Dialogue Tasks," 2018, https://arxiv.org/abs/1811.11707, last visit July 7, 2020.

[33] A. Xu, Z. Liu, Y. Guo, V. Sinha, and R. Akkiraju, "A New Chatbot for Customer Service on Social Media," in *ACM CHI Conference on Human Factors in Computing Systems*, Denver, CO, USA, May 2017, pp. 3506–3510.

[34] V. Yadav and S. Bethar, "A Survey on Recent Advances in Named Entity Recognition from Deep Learning model," in *International Conference on Computational Linguistic (COLING)*. Santa Fe, New Mexico, USA: Association for Computational Linguistic, August 2018, pp. 2145–215.

[35] O. Yavanoglu and M. Aydos, "A Review on Cyber Security Datasets for Machine Learning Algorithms," in *IEEE International Conference on Big Data (Big Data)*, Boston,MA,USA, June 2017, pp. 2186–2193.

[36] P. Zambrano, M. Sanchez, J. Torres, and W. Fuertes, "BotHook: An Option Against Cyberpedophilia," in *1st Cyber Security in Networking Conference (CSNet)*, Rio de Janeiro, Brazil, October 2017, pp. 1–3.

[37] L. Zhou, J. Gao, D. Li, and H.-Y. Shum, "The Design and Implementation of Xiaoice, an Empathetic Social Chatbot," *Computational Linguistics*, vol. 46, no. 1, pp. 53–93, 2020.